Personality and Individual Differences 219 (2024) 112514

Contents lists available at ScienceDirect

Personality and Individual Differences

journal homepage: www.elsevier.com/locate/paid

# Modeling personality language use with small semantic vector subspaces

José Ángel Martínez-Huertas [a,*], Guillermo Jorge-Botana [b], Alejandro Martínez-Mingo [c], José David Moreno [c], Ricardo Olmos [c]

[a] Faculty of Psychology, National Distance Education University, Spain
[b] Faculty of Psychology, Complutense University of Madrid, Spain
[c] Faculty of Psychology, Autonomous University of Madrid, Spain

## 1. Introduction

Language has played a crucial role in the development of personality theory and has profoundly influenced the study of individual differences. For example, the structure of personality has been constructed through descriptions and factor analysis under some paradigms, like the well-known *lexical hypothesis* of personality (e.g., Ashton & Lee, 2005; see also the excellent revision by Goldberg, 1993). But it was the seminal study of Pennebaker and Graybeal's (2001) which opened the door to find relevant indicators of personality traits in written language by means of human experts' criteria and computational assessments. Nowadays, neural networks, specially RNN-LSTMs and Transformers, have proven to be efficient tools in analyzing text output, such as utterances, and have been used to predict mood states or discourse types in sentences and other language productions. Nonetheless, neural networks are constrained by the number of samples required for training the models (i.e., they need a big set of utterances to reach a useful model). This is a handicap for the development of computational modeling of personality language use (PLU) that, in our opinion, could be eased by using specific semantic vector subspaces. In other words, we believe that using information from primitive semantic indicators could reduce the costs of obtaining indicators of personality traits from language. In this sense, the semantic indicators of PLU can be seen as a set of resumes of observable information that makes language properties more manageable and systematic. However, this philosophy requires the formalization of such semantic properties of each personality trait.

There are some proposals that have formalized and collected language primitives. For example, the well-known LIWC indicators from Pennebaker et al. (2001) serve as a paradigmatic example of how different observable indicators (a set of lexical and grammatical clues) are used for making predictions in a possible subsequent model. Multiple articles have been published remarking different interesting relationships between personality and written language under the assumption that there are characteristics of personality which are embedded and reflected in the patterns of language that people use (e.g., Boyd & Pennebaker, 2017; Fast & Funder, 2008; Moreno et al., 2021; Pennebaker & King, 1999). However, these approaches have two main limitations. First, the language-personality relations do not process abstract information as it just counts primitives. Second, the semantic cues do not accurately generalize the relationship with personality, as they are based solely on literal occurrences of words from a general dictionary. Thus, the personality relevant open-class terms are underrepresented in these dictionaries, as well as their possible connotations (Garten et al., 2018).

The basic idea of the present proposal is the generation of a mapping that summarizes language information by a priori semantic primitives defined by the researcher. For this purpose, the dimensions of vector space models (VSMs) are highly suitable and cost-effective tools to serve as such indicators. Semantic vectors have components which represents an exhaustive representation of the semantics as the score in one dimension indicates the presence of a certain meaning (a topic) in a text. These topics are not based in literal occurrences like the LIWC indicators (e.g., Garten et al., 2018). Conversely, topics from VSMs are based on automatic processing of a large sample of text, where all the words are represented with vectors. The dimensions of those vectors are identified to maximize the comprehension and representation of the semantics (e.g., Günther et al., 2019; Jorge-Botana et al., 2020; or McNamara, 2011; see the next section for an introduction on VSMs for the study of personality traits). VSMs have been found to capture deeper or not apparent semantic characteristics of language.

But the indicators provided by the vector of an ordinary VSM still lacks specificity. The optimum scenario is to have a mapping between the primitives of PLU with indicators of semantic relevant properties. New developments from VSMs, like the generation of semantic vector

* Corresponding author at: Department of Methodology of Behavioral Sciences, National Distance Education University, C. de Juan del Rosal, 10, 28040 Madrid, Spain.

*E-mail addresses:* jamartinez@psi.uned.es (J.Á. Martínez-Huertas), guijorge@ucm.es (G. Jorge-Botana), josedavid.moreno@uam.es (J.D. Moreno), ricardo.olmos@uam.es (R. Olmos).

https://doi.org/10.1016/j.paid.2023.112514
Received 25 July 2023; Received in revised form 14 October 2023; Accepted 4 December 2023

subspaces, are a promising perspective to formalize indicators with only relevant properties. A subspace is a less dimensional vector space, inside an ordinary and general domain VSM space, which focuses on specific information in some selected dimensions and whose subspace coordinates identify a word or a text in some selected semantic properties. Technically speaking, it is spanned only with a set of selected vectors (a basis) relevant to a concept, in our case, a personality trait (see a similar rationale, based on word norms produced by participants, in Kjell et al., 2019). Therefore, the properties of PLU of utterances can be captured just projecting those utterances into the subspace. In other words, it is possible to express those utterances with the references of the subspace. As a result, these subspaces are expected to be markers sensible to the presence of the semantic specific to a trait. At the end, these indicators could even allow to have a massive set of observable clues to be included in other predictive models applied to identify different personality styles or coping styles from simple written productions (constructed responses, letters, or scenario-based essays) or written productions from social networks (Twitter, Facebook, or blogs).

In this paper, we aim to illustrate this *proof of concept* as a tentative way of generating low-cost indicators that summarize observable occurrences of words into observable semantic indicators with the goal of capturing relevant variability of personality traits from language. This is done in an effort to complement and enhance other systems, such as the LIWC or more sophisticated predictive models like RNN-LSTMs or Transformers with embedding layers. To generate the indicators of PLU, we are going to use vector subspaces generated with a procedure proposed in previous studies (Jorge-Botana et al., 2019; Martínez-Huertas et al., 2022). This procedure emulates a bifactor structure, allowing us to measure general and specific semantic meanings relevant to a concept. As we will see, these semantic vector subspaces are like a room whose walls are made of semantic relevant properties of a personality trait. The location in these rooms will identify an utterance in the trait, and they could be used to extract personality trait-relevant semantic properties from words in utterances. Specifically, we will create indicators based on the Big Five personality traits, namely: openness to experience, conscientiousness, extraversion, agreeableness, and neuroticism (McCrae & Costa Jr., 2008). However, this proposal is generalizable to other models of personality or individual differences. Summarizing, our proposal is to create five semantic vector subspaces to measure general and specific semantic meanings related to the Big Five model of personality to obtain low-cost pseudo-observable indicators of personality from words in people's utterances. To do so, we evaluated our participants in different prompted-based self-descriptions to analyze the generalizability of the language indicators in our empirical illustration. In this proof of concept, different standardized self-report inventories used to measure the Big Five personality traits were considered as validity criteria of our language indicators.

### 1.1. *Generating indicators of personality language use (PLU) in semantic vector subspaces*

VSMs algebraically vectorize word occurrences to represent the lexicon in a reduced dimensionality vector space (see Günther et al., 2019; Jorge-Botana et al., 2020; or McNamara, 2011 for a revision on VSMs). This allows for the creation of a more abstract semantic layer of meaning representation, rather than relying solely on the information from literal occurrences of words in texts. The dimensions (the vectors of the semantic space basis) produced by most VSMs are latent and have no explicit meanings. They are useful and parsimonious, but it is difficult to identify their meanings. For this reason, there have been efforts to generate interpretable dimensions in VSMs from both exploratory (Kundu et al., 2015; Visinescu & Evangelopoulos, 2014) and confirmatory (Hu et al., 2007; Jorge-Botana et al., 2019; Martínez-Huertas et al., 2021; Martínez-Huertas et al., 2022) approaches. In the present paper, we are going to work with a procedure that has been validated in the well-known latent semantic analysis or LSA model (e.g., Landauer & Dumais, 1997).

Some of these techniques aiming to generate interpretable dimensions have been used to detect the presence of semantic concepts in academic texts, where each concept is treated as a subspace with some relevant and meaningful dimensions. For example, if a student's text scores in such relevant semantic dimensions, it can be said that the text covers the target concept. For example, the presence of concept "Theory of evolution" can be estimated by scores in relevant dimensions like "phenotypes and gene expression", and/or "natural selection". The procedure of generation of such subspaces has been called Inbuilt Rubric for people working in educational assessment as it emulates an assessment rubric (e.g., Jorge-Botana et al., 2019; Martínez-Huertas et al., 2021, 2022).

In this line, considering a concept in a generic sense, a personality trait can be also considered a concept that can be defined by the presence of a set of references, that is, a set of scores in some dimensions. For instance, the presence of the trait of extraversion can be assessed by evaluating scores on relevant dimensions such as "social interactions" or "participation in group activities". In addition, the set of relevant dimensions is a kind of indicators wherein the presence or absence of certain key topics can be discerned. In the context of personality, there have been attempts of capturing Big Five semantic properties using personality trait definitions in VSMs (Kwantes et al., 2016). However, they heavily relied on cosine measures to evaluate the similarity between vectors, which is a scalar measure with limitations when compared to multidimensional semantic representations (Martínez-Huertas et al., 2021).

We propose to formalize each personality trait of the Big Five model (McCrae & Costa Jr., 2008) with subspaces of two dimensions (a direct and an inverse language indicator) plus an additional general dimension based on hierarchical VSMs (Jorge-Botana et al., 2019; Martínez-Huertas et al., 2022) in the LSA model (Landauer & Dumais, 1997). In order to align our proposal with existing psychometric models of the Big Five personality traits, we used the items of the *Ten Item Personality Measure* (TIPI; Gosling et al., 2003) to operationalize the definition of the personality traits. It was an affordable tool for this task due to it only has two items (one direct and one inverse item) for each trait. In any case, it is worth mentioning that there are other options to formalize PLU, although we preferred to use one of the simplest definitions of these personality traits for this proof of concept.

The procedure to construct a subspace with those dimensions begins with the collection of a sample of words that represents the direct and inverse indicators. For example, the words of the TIPI items could be collected as the direct indicator ("*open to new experiences*, *multi-faceted*") and as the inverse one ("*traditional*, *unimaginative*"). For both set of words, a vector is calculated summing the vectors of each isolated word as follows:

$$v_{direct} = v_{open} + v_{new\_experiences} + v_{multifaceted}$$

$$v_{inverse} = v_{traditional} + v_{unimaginative} \quad (1)$$

Thus, $v_{direct}$ and $v_{inverse}$ are two vectors that represent part of the two poles of the personality trait. To construct a new subspace that includes these vectors as part of its basis, a change of basis is required to convert the original latent vector space into a new space. Let us assume that the basis of the original latent space is represented in a matrix $B$ resulting from SVD (the LSA procedure to reduce dimensionality). A possible transformation matrix is a matrix $\boldsymbol{\beta}$ with the same number of columns of the matrix $B$ but having the two first columns filled with the vectors $v_{direct}$ and $v_{inverse}$. We need also a third vector that represents a general dimension of $v_{direct}$ and $v_{inverse}$, that encompasses the common variance between the direct and inverse semantic indicators.[1] To identify that vector, we use the *split* method proposed in Jorge-Botana et al. (2019). This method constructs a matrix with all the vectors of the words in the direct and inverse descriptors and run an Exploratory Factor Analysis. Then, the general dimension ($v_{general}$) is obtained by weighting the semantic vectors by their respective factor loadings ($\lambda_{ki}$) like this:

$$v_{general} = \lambda_{1i} * Direct_{1i} + \lambda_{2i} * Inverse_{2i} + error_{ki} \quad (2)$$

where $k$ is the number of language indicators ($k = 2$), and $i$ is the number of the partitions of the lexical descriptors of each concept (in the *split* method parametrization: $i = 1, 2$). The role of the general dimensions of the hierarchical vector subspace is twofold: (1) they aim to capture actual and relevant variability of the target concepts, and (2) they are useful to distill the semantic representations of the specific vectors which, in this case, correspond to the direct and indirect language indicators (see Jorge-Botana et al., 2019 and Martínez-Huertas et al., 2022 for a discussion about this topic in the context of the evaluation of summaries from instructional texts in educational settings).

At this point, the first three columns of **β** are filled with $v_{direct}$, $v_{inverse}$ and $v_{general}$. Since the transformation matrix require the same matrix dimensionality as $B$, such three-dimensional **β** matrix is randomly filled with vectors from the standard basis:

$$\beta = \left( v_{direct} \; v_{inverse} \; v_{general} \; v_{standard\ basis\ 1} \cdots v_{standard\ basis\ n} \right) \quad (3)$$

This **β** matrix allow us to change the basis of the original vector space $U$ into a new vector space $U'$ whose three first dimensions mean the direct, inverse and general indicators, respectively. This is the operation:

$$U' = U\beta^T \quad (4)$$

Resulting in a new vector space $U'$ where each word can be described with a vector whose three first dimensions are relevant indicators for the target personality trait. To avoid the distortion of the original semantic distances of the words in the vector space, the vectors of matrix **β** are orthogonalized by the Gram–Schmidt procedure, which is repeated multiple times in a random order to avoid potential order-bias. The Gram–Schmidt procedure converts the change of basis in a simple rotation of $U$ into $U'$. Also, a correlation (e.g., $Cor(v_{direct}, v'_{direct})$) above of 0.85 is imposed between the vector before and after rotation to ensure that the meaning of the vector is not significantly different to the previous vector.

Once the new vector space $U'$ has been generated, it is possible to represent any utterance $d$ as a vector $d'$ in the new space by a simple projection (authors usually call this projection as *folding-in* and $d'$ as *pseudodocument*, see Landauer & Dumais, 1997 for more details).

$$d' = d^T U' \quad (5)$$

where the first three dimensions of vector $d'$ would capture the quantity of language related to the semantic meanings covered by $v_{direct}$, $v_{inverse}$ and $v_{general}$. See Jorge-Botana et al. (2019) and Martínez-Huertas et al. (2022) for additional details on the methodology used to generate these hierarchical VSMs.

It can be said that the final product of the procedure is like a bifactor modeling structure because of the specific and the general language indicators do not share common variance (they are orthogonal dimensions). Fig. 1 graphically represents the semantic vector subspace of openness and the position of two individuals (*S1*: triangle, *S2*: circle) based on their coordinates in the direct (*O1*), inverse (*O2*), and general (*OG*) language dimensions of their responses. This would be the result of projecting the answers of these individuals using Eq. (5). In this example, the individual S2 would present larger coordinates in the dimensions of the subspace than the individual S1 (which would have more common or typical language). Different patterns of PLU could be found depending on the low/medium/high scores of the direct, the inverse, and the general dimensions of the semantic vector subspace.

*Note*. Three orthogonal dimensions configurate this low-cost indicator of personality language use (PLU) for openness: (1) open to new experiences; multifaceted (direct language indicator; O1), (2) traditional; unimaginative (inverse language indicator; O2); and (3) the general dimension (OG).

### 1.2. *Objective*

The aim of the present paper is to test a proof of concept, namely: a potential method for generating low-cost indicators of language use that capture relevant variability of personality traits from language through semantic vector subspaces. Specifically, these semantic vector subspaces, which are unique to each personality trait, represent a proposal for formalizing the semantics of the PLU (in this case, of the Big Five personality model, e.g., McCrae & Costa Jr., 2008). As previously discussed, we will measure general and specific semantic meanings relevant to the definition of each personality trait, using recent computational developments in VSMs (e.g., Jorge-Botana et al., 2019; Martínez-Huertas et al., 2022). The resulting indicators of PLU generated through semantic subspaces will be evaluated in two distinct prompted-based self-descriptions (i.e., study 1 and study 2). In the first study, we are going to analyze the convergent and discriminant validity of different latent profiles of PLU. These are discrete latent variables resulting from latent profile analysis that aim to characterize different patterns of PLU. We expect to find a large profile of people using common language (that is, without low nor high scores of the language indicators) and different profiles of PLU patterns. In the second study, we will analyze the linear relations between the language indicators and the personality traits, highlighting the differences between the profiles found in the first study. We expect to find larger language-personality relations in profiles that exhibit relevant personality topic use in comparison to the profile of individuals using common language. Additionally, different parametrizations of the semantic vector subspaces will be analyzed to model language use of personality traits of the Big Five (that is, considering both classic and hierarchical VSMs, Jorge-Botana et al., 2019; Martínez-Huertas et al., 2022). In this sense, we are also going to test if the general dimensions of this computational method are able to capture relevant variance of the personality traits, or if they are just a way of distilling the other vector subspace dimensions (see Martínez-Huertas et al., 2022 for a full discussion about this point). Finally, we would like to remark that the present study could serve as an illustration of the use of semantic vector subspaces to formalize the semantics of language to ease the Big Five personality screening in texts, but this methodology would be applicable to the study of other individual differences.

Methods and materials (studies 1 and 2).

### 1.3. *Participants*

A total of 643 subjects participated in this study. They were undergraduate Spanish students who received extra course credits for their participation in the study. Their average age was 19.50 years ($SD = 2.22$; 13 % males). All of them were native Spanish speakers.

### 1.4. *Instruments*

#### 1.4.1. *Standardized personality assessments*

Two different self-report inventories with Likert-type answer scales were used to measure the Big Five personality traits:

[1] The computation of the general dimension ($v_{\text{general}}$) is optional. In that case, the semantic subspace would be computed with only two dimensions ($v_{\text{direct}}$ and $v_{\text{inverse}}$). This perspective, which would result in two-dimensional semantic vector subspaces, was called *classic* in this paper.

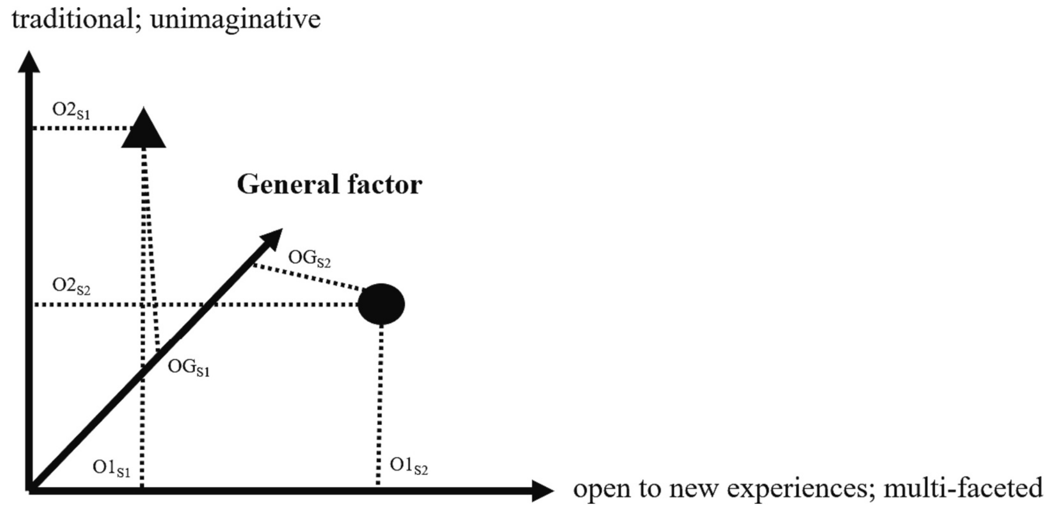


**Fig. 1.** Graphical representation of the coordinates of two individuals (S1: triangle, S2: circle) in a three-dimensional semantic vector subspace for openness personality trait.

1. The Spanish version of the Big Five questionnaire (Bermúdez, 2001). Appropriate reliability was found for the personality traits: openness (O; Cronbach's $\alpha = 0.79$; McDonald's $\omega = 0.83$), conscientiousness (C; $\alpha = 0.86$; $\omega = 0.89$), extraversion (E; $\alpha = 0.75$; $\omega = 0.79$), agreeableness (A; $\alpha = 0.82$; $\omega = 0.85$), and neuroticism (N; $\alpha = 0.90$; $\omega = 0.92$).
2. The Spanish version of the OPERAS questionnaire (Vigil-Colet et al., 2013). Reliability was good for all the personality traits: O ($\alpha = 0.78$; $\omega = 0.80$), C ($\alpha = 0.73$; $\omega = 0.75$), E ($\alpha = 0.87$; $\omega = 0.87$), A ($\alpha = 0.70$; $\omega = 0.72$), and N ($\alpha = 0.84$; $\omega = 0.85$).

Table 1 presents the descriptive analysis and the Pearson correlation coefficients of the personality traits of both questionnaires. The mean Pearson correlation coefficient for all the personality traits was 0.698 ($r^2 = 0.487$), which represents the mean relationship between the same personality construct measured with different questionnaires. A mean score was obtained for each personality trait using both questionnaires. There are medium correlation coefficients between the scores of the different personality traits, which would be a handicap for the performance of the computational scores of the language-based tasks.

### 1.4.2. *Constructed responses (language-based tasks)*

Supplementary materials present the prompts and the automatically translated responses of some participants as an example of the answers in each task.

1. 10-word self-descriptions: Participants were instructed to compose a succinct self-description of ten words. On average, participants submitted 9.90 words ($SD = 0.63$, min = 4, max = 11).
2. Self-descriptions for a social profile: Participants were tasked to produce a self-description of approximately 500 words. On average, participants wrote 363.37 words ($SD = 153.03$, $min = 41$, $max = 682$).

### 1.4.3. *Computational measures (evaluation of language-based tasks)*

Different materials and software were used to compute the evaluation of language-based tasks:

1. Firstly, a 300-dimensional vector space was generated using the standard *Latent Semantic Analysis* procedure (Landauer & Dumais, 1997) via *GallitoStudio* software (Jorge-Botana et al., 2013). A random sample of 455,969 documents (paragraphs) from a random sample of the Spanish Wikipedia served as the linguistic corpus for this vector space generation using log-entropy as the weighted function. A total of 70,244 unique terms were processed to generate the original VSM. The resulting semantic vector representations were normalized. This corpus has been validated in previous studies with a special focus on historical and scientific/technological concepts (e.g., Martínez-Huertas et al., 2022).
2. Secondly, different subspaces were generated for each of the personality traits with direct and inverse dimensions. The classic and hierarchical versions of the Inbuilt Rubric method (with or without adding the vector of the general dimension to the basis) were applied

**Table 1**
Means (*M*), standard deviations (*SD*), and Pearson correlation coefficients of the Big Five personality traits for BFQ and OPERAS standardized self-report inventories.

| | $O_1$ | $C_1$ | $E_1$ | $A_1$ | $N_1$ | $O_2$ | $C_2$ | $E_2$ | $A_2$ | $N_2$ |
|---|---|---|---|---|---|---|---|---|---|---|
| $O_1$ | - | .27** | .37** | .44** | .16** | .65** | .23** | .24** | .33** | .21** |
| $C_1$ | | - | .28** | .22** | .00 | .24** | .74** | .08 | .13** | .11** |
| $E_1$ | | | - | .18** | -.01 | .20** | .33** | .71** | -.03 | .26** |
| $A_1$ | | | | - | .22** | .40** | .24** | .26** | .71** | .24** |
| $N_1$ | | | | | - | .04 | .19** | .02 | .26** | .68** |
| $O_2$ | | | | | | - | .17** | .10** | .32** | .09* |
| $C_2$ | | | | | | | - | .18** | .18** | .31** |
| $E_2$ | | | | | | | | - | .01 | .30** |
| $A_2$ | | | | | | | | | - | .20** |
| $N_2$ | | | | | | | | | | - |
| *M* | 3.60 | 3.63 | 3.15 | 3.63 | 2.59 | 4.12 | 3.56 | 3.22 | 3.83 | 3.08 |
| *SD* | .42 | .47 | .42 | .41 | .61 | .61 | .63 | .82 | .52 | .81 |

(Jorge-Botana et al., 2019; Martínez-Huertas et al., 2021, 2022). The procedure was explained in previous sections. This computational method turns the semantic vector space into a new semantic space where the two or three first dimensions are personality trait relevant. Descriptors for constructing the direct and inverse vectors were based on the items of the *Ten Item Personality Measure* (TIPI; Gosling et al., 2003; see the Procedure section). The transformation of the semantic vectors from the original vector space to the new semantic vector subspace was reliable (i.e., no semantic distortions were observed).
3. Once the semantic vector subspaces of each personality trait were generated, participants' constructed responses to the language-based tasks were projected into them. As explained in the Introduction section, only two/three coordinates of the semantic vector subspace were used to study their relations with the scores of the standardized self-report inventories.

For a more detailed explanation of the computational method, please refer to the Introduction and Procedure sections.

### *1.5. Procedure*

Firstly, participants were tasked to produce a self-description in approximately ten words. Secondly, they made self-descriptions for a social profile in, approximately, 500 words. Additionally, participants completed two standardized self-report inventories for evaluating the Big Five personality traits. To analyze their responses, a 300-dimensional vector space was generated using the standard *Latent Semantic Analysis* procedure (e.g., Landauer & Dumais, 1997) with the Spanish Wikipedia as the corpus. Then, the Inbuilt Rubric method, in both its classic and hierarchical versions (Jorge-Botana et al., 2019; Martínez-Huertas et al., 2021, 2022), was applied to generate new semantic vector subspaces for each of the Big Five personality traits. To construct the direct and indirect vector indicators of each basis, we used the items of the Spanish version of another standardized self-report inventory, the TIPI (Gosling et al., 2003). Five different semantic vector subspaces, one for each trait, were produced with the following descriptors:

1. Openness:
   a. open to new experiences; multi-faceted
   b. traditional; unimaginative
2. Conscientiousness:
   a. reliable; self-disciplined
   b. disorganized; careless
3. Extraversion:
   a. enthusiastic; extravert
   b. quiet; reserved
4. Agreeableness:
   a. choleric; argumentative
   b. understanding; kind
5. Neuroticism:
   a. anxious; easily upset
   b. serene; emotionally stable

The specific procedures and analyses of each study are described in the following sections. All the words used in this study were lemmatized. The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of Autonomous University of Madrid.

## 2. Study 1: estimation and validation of profiles of personality language use (PLU)

In the first study, we estimated and validated different profiles of PLU from the 10-word self-descriptions. Once the constructed responses were evaluated using the semantic vector subspaces, a latent profile analysis was conducted in *Mplus 8* (Muthén & Muthén, 1998-2017). In the semantic vector subspace generated with the original Inbuilt Rubric method, the direct and inverse language use dimensions were used as the input of the latent profile analysis. In the hierarchical semantic vector subspace (with general dimension), the analyses were conducted using two different sets of computational scores as input: using the direct and inverse language use dimensions, and the general language use dimensions. The latent profile analysis was conducted freely estimating the means and variances of the PLU indicators. The number of latent profiles of each personality trait was selected based on the model fit indices with the requirement of presenting enough participants in all the latent profiles ($n \geq 20$). Then, the convergent and discriminant validity of the latent profiles of PLU was analyzed using logistic regressions of the Mplus' *3-STEP* procedure (Muthén & Muthén, 1998-2017) using the scores of the standardized self-report inventories as criteria. That is, we analyzed the relations between the self-report inventories (covariates) and the latent profiles (dependent variables). Such convergent and discriminant validity was considered adequate when the standardized self-report inventories presented a statistically significant relationship with the language use latent profiles of the target trait, and a null or less important relationship with the rest of the traits.

### *2.1. Describing computational scores of personality language use (PLU)*

All the computational variables were standardized. Thus, those computational scores around zero represent a common language use according to that PLU indicator, indicating neither a significant presence nor absence of prototypical language use for that personality trait. Those computational scores with larger values represent a more prototypical language use for that indicator (e.g., in openness, they represent prototypical language of open people or not-open people). Fig. 2 presents different scatterplots to illustrate the variability of the PLU indicators in the two-dimensional semantic vector subspaces (similar pattern of results were observed for the dimensions of the hierarchical semantic vector subspaces). These figures reveal distinct patterns of PLU across the different personality traits. In any case, the linear relationships between the direct and inverse language indicators were null for openness, small for conscientiousness, extraversion and neuroticism, and medium for agreeableness.

### *2.2. Defining the latent profiles of personality language use (PLU)*

An exhaustive study of the latent profiles of PLU was conducted for (1) the dimensions of the two-dimensional semantic vector subspaces, (2) the specific dimensions of the hierarchical semantic vector subspaces, and (3) the general dimensions of the hierarchical semantic vector subspaces. Table 2 presents the model fit results for the selection of the number of latent profiles in each personality trait (for the sake of brevity, we only present the results of the best performing version of the semantic vector subspaces for each personality trait). This was done to explore all the natural combinations of PLU indicators. As it can be seen, the hierarchical semantic vector subspaces were selected as the most appropriate for extracting the latent profiles of PLU in four out of five traits.

Table 3 and Fig. 3 present the means of the indicators of PLU of each profile in each personality trait. There are different profiles of PLU underlying the computational indicators. In the case of openness, the first latent profile was characterized for a medium level of language use for both indicators, whereas the second one exhibits a larger use of language related to the direct indicator. The same pattern of language use was found in the latent profiles of neuroticism. In the case of conscientiousness, the first latent profile was characterized by a medium level of language use for both indicators, but the second one was characterized by a larger use of language related with the direct indicator and a lower use of language associated with the inverse indicator. In agreeableness, a lower use of direct and inverse indicators was found for the first latent

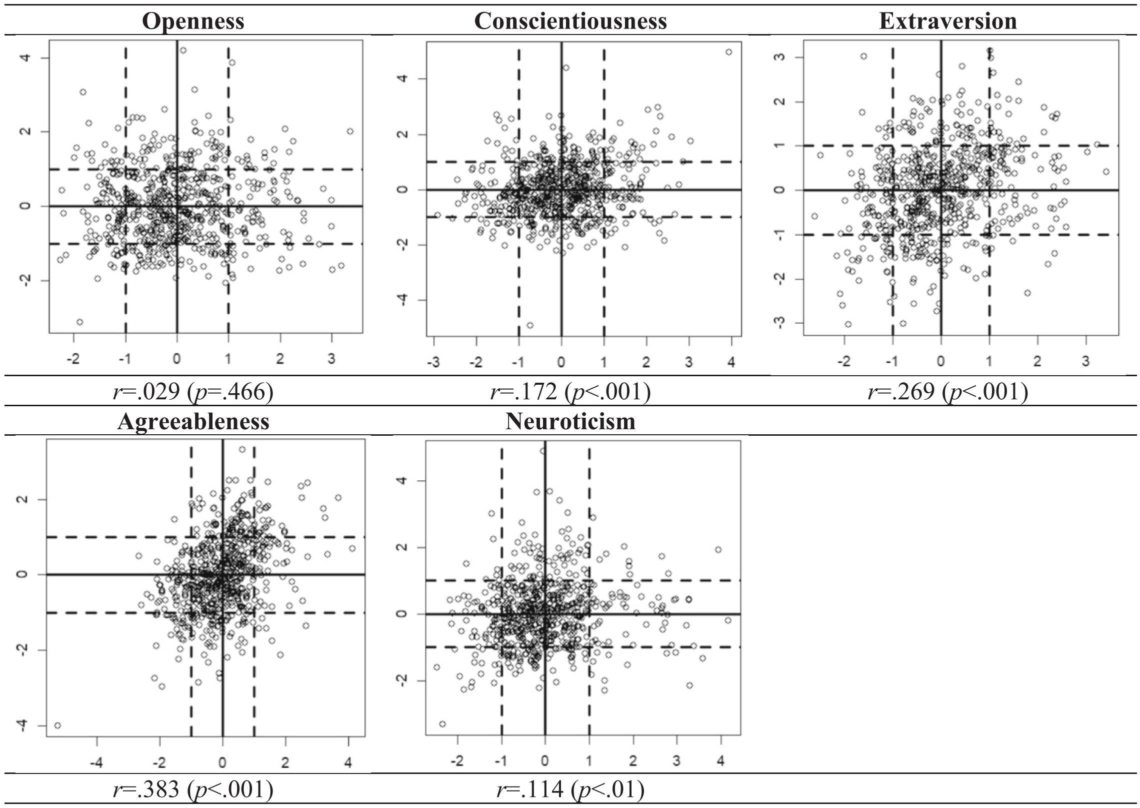


**Fig. 2.** Direct and inverse language indicators in the two-dimensional semantic vector subspaces (10-word self-descriptions).
*Note*. x-axis: direct indicator. y-axis: inverse indicator. Continuous lines: means. Discontinuous lines: ±1 standard deviation.

**Table 2**
Model fit results for models with different number of latent profiles in each personality trait.

| Personality trait | Sem. vector subspace | Language indicators | N. latent profiles | Fit indices | | | | |
|---|---|---|---|---|---|---|---|---|
| | | | | Loglikelihood | AIC | BIC | saBIC | Entropy |
| Openness | H | DI | 1 | −1815.241 | 3638.481 | 3656.327 | 3643.627 | – |
| | | | **2** | **−1798.386** | **3610.771** | **3642.002** | **3619.777** | **0.707** |
| | | | 3 [a] | −1795.842 | 3611.684 | 3656.299 | 3624.549 | 0.565 |
| Conscientiousness | T | DI | 1 | −1815.241 | 3638.481 | 3656.327 | 3643.627 | – |
| | | | **2** | **−1713.670** | **3441.340** | **3472.570** | **3450.346** | **0.904** |
| | | | 3 [a] | −1713.670 | 3447.340 | 3491.955 | 3460.205 | 0.397 |
| Extraversion | H | G | 1 | −907.620 | 1819.241 | 1828.163 | 1821.814 | – |
| | | | **2** | **−906.107** | **1820.214** | **1838.060** | **1825.360** | **0.555** |
| | | | 3 [a] | −900.803 | 1813.607 | 1840.376 | 1821.326 | 0.702 |
| Agreeableness | H | DI | 1 | −1815.241 | 3638.481 | 3656.327 | 3643.627 | – |
| | | | **2** | **−1752.276** | **3518.551** | **3549.781** | **3527.557** | **0.566** |
| | | | 3 [a] | −1733.491 | 3486.982 | 3531.597 | 3499.847 | 0.750 |
| Neuroticism | H | DI | 1 | −1815.241 | 3638.481 | 3656.327 | 3643.627 | – |
| | | | **2** | **−1786.444** | **3586.887** | **3618.117** | **3595.893** | **0.903** |
| | | | 3 [a] | −1767.107 | 3554.215 | 3598.829 | 3567.080 | 0.859 |

*Note*. Bold/highlight = Selected model. T = Two-dimensional semantic vector subspace. H = Hierarchical semantic vector subspace. DI = Direct and inverse indicators. G = General dimension. a = Model was not-selected due to a small number of participants in one or different latent profiles.

profile, and a larger use of both indicators in the second latent profile. This could be evidencing that the model was not able to differentiate between the semantic contents used to generate the semantic vector subspace in the agreeableness trait. Lastly, the extraversion latent profiles, based on the general dimension, presented a first latent profile with lower language use related to the general dimension, while the second latent profile was characterized by a medium use of extraversion language.

*Note*. x-axis = Latent profiles. 95 % confidence intervals were computed using the standard error of the estimated means.

### *2.3. Convergent and discriminant validity of the profiles of personality language use (PLU)*

As stated previously, we expect to find convergent and discriminant validity for the latent profiles of PLU regarding to the scores of the standardized self-report inventories. Table 4 presents the logistic regression results where different covariates (standardized self-report inventories) were used to classify the latent profiles of PLU (P1-P2; dependent variable) with Mplus *3-STEP* procedure. Appropriate convergent validity was found for the resulting latent profiles of all the personality traits. Moreover, given that the variables were standardized and that the standard deviation of the logit scale is approximately equal

**Table 3**
Means and variances of language indicators of the latent profiles of each personality trait.

| Personality trait | Language indicator | P1 | | | P2 | | |
|---|---|---|---|---|---|---|---|
| | | N | Mean (*SE*) | Var (*SE*) | N | Mean (*SE*) | Var (*SE*) |
| O | Direct | 554 | −0.282*** (0.054) | 0.587*** (0.045) | 86 | 1.458*** (0.159) | 0.587*** (0.045) |
| | Inverse | | −0.020 (0.047) | 0.996*** (0.060) | | 0.102 (0.142) | 0.996*** (0.060) |
| C | Direct | 575 | −0.240** (0.077) | . 0.492** (0.172) | 65 | 2.107*** (. 510) | 0.492** (0.172) |
| | Inverse | | −0.121* (0.061) | . 870*** (0.074) | | −1.061*** (0.179) | 0.870*** (0.074) |
| E | General | 48 | −1.089*** (0.228) | 0.786*** (0.156) | 592 | 0.195 (0.124) | 0.786*** (0.156) |
| A | Inverse | 408 | −0.448*** (0.069) | 0.668*** (0.055) | 232 | 0.738*** (0.085) | 0.668*** (0.055) |
| | Direct | | −0.496*** (0.062) | 0.594*** (0.054) | | 0.816*** (0.106) | 0.594*** (0.054) |
| N | Direct | 605 | −0.154*** (0.038) | 0.652*** (0.043) | 35 | 2.257*** (0.225) | 0.652*** (0.043) |
| | Inverse | | −0.015 (0.043) | 0.995*** (0.072) | | 0.224 (0.221) | 0.995*** (0.072) |

*Note.* *** = $p < .001$. ** = $p < .01$. * = $p < .05$. P1-P2 = Latent profiles 1 to 2. O = Openness. C = Conscientiousness. E = Extraversion. A = Agreeableness. N = Neuroticism.

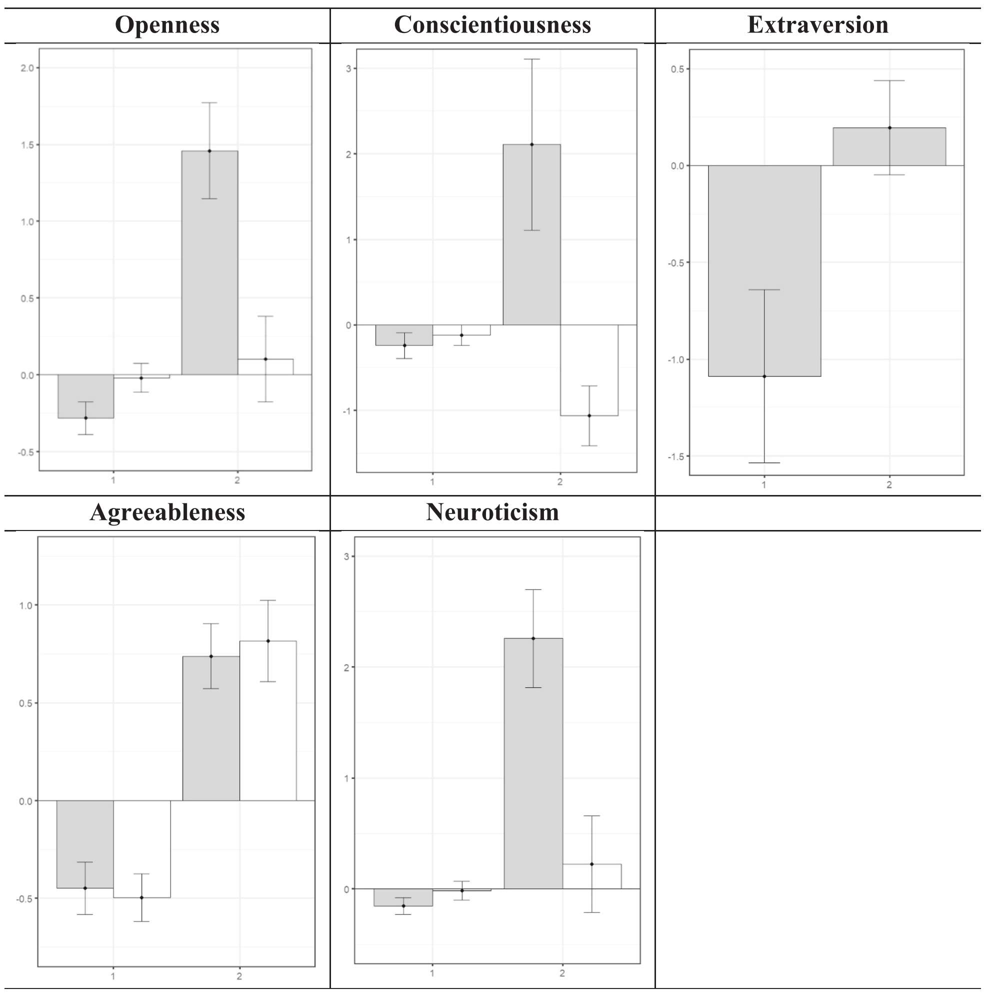


**Fig. 3.** Means of the direct and inverse indicators of personality language use (PLU) of each profile in each personality trait.

to 1.81, the effect sizes of the estimated logistic regression coefficients can be considered moderate for all the latent profiles, except for agreeableness that would be small and for extraversion that would be large. Additionally, we can see that higher scores in the personality traits measured with the standardized self-report inventories predict larger probability of being classified in the PLU profiles with relevant means in the language indicators of its trait.

## 3. Study 2: further validation of personality language use (PLU)

In the second study, further validation evidence was found in favor of PLU profiles and the computational indicators of the semantic vector subspaces using the self-descriptions for a social profile. These self-descriptions were significantly longer than the 10-word self-descriptions, roughly 30 times. Since personality has been considered a hierarchical construct, it is worth investigating if different levels of the construct could lead to differences in language-personality relations (e.g., Digman, 1997; Eysenck, 1994; McCrae & Sutin, 2018). Such differential results on the relations between a variable and a hierarchical construct with different levels (like personality) has been referred to as the Brunswik symmetry principle (e.g., Nesselroade & McArdle, 1997; Wittmann, 1988). As a representative example, Nesselroade and

**Table 4**
Logistic regression coefficient ($\beta$) results (Mplus *3-STEP*) to analyze convergent and discriminant validity of latent profiles of personality language use (PLU) by standardized self-report inventories' scores.

| | | Latent profiles of PLU | | | | |
|---|---|---|---|---|---|---|
| | | O | C | E | A | N |
| Standardized self-report inventories | O | .996** (.289) | -.468* (.233) | -.249 (.188) | -.292 (.236) | .094 (.191) |
| | C | -.123 (.191) | .961*** (.218) | .240 (.178) | -.272 (.233) | .206 (.192) |
| | E | .359 (.400) | -.676 (.710) | -2.005** (.634) | .520 (.539) | -.598 (.414) |
| | A | -.262 (.166) | -.070 (.144) | -.055 (.147) | .603*** (.165) | -.319* (.144) |
| | N | -.420 (.344) | -.328 (.321) | .174 (.278) | .385 (.273) | 1.295*** (.362) |

*Note.* *** = $p < .001$. ** = $p < .01$. * = $p < .05$. Grey shading = Appropriate convergent and discriminant validity. O = Openness. C = Conscientiousness. E = Extraversion. A = Agreeableness. N = Neuroticism. Reference of logistic regression coefficients = Latent profiles numbered as 1.

McArdle (1997) concluded that two empirical correlations could be underestimating the true correlation between two constructs when there is a mismatch between their hierarchical levels of abstraction in the mathematical-statistical model.

In this section, we are going to study the relationships between the PLU indicators from semantic vector subspaces and both dimensions and facets of the Big Five questionnaire (Bermúdez, 2001) to evaluate the aforementioned Brunswik symmetry principle. To do so, we employed multiple-group structural equation models, using the R's lavaan package (Rosseel, 2011) with diagonally weighted least squares estimator, to examine the associations between the language indicators and the dimensions/facets of the questionnaire. These results were tested in a multi-group analysis, using the PLU profiles of study 1, for each of the personality traits, to evaluate if the language-personality relations were similar in both profiles of PLU. We chose to conduct separate analyses for each personality trait, as including the computational scores of all traits resulted in convergence problems in the models, a phenomenon that has been previously observed in computational assessments due to their generation from orthogonal dimensions (Martínez-Huertas et al., 2022).

### 3.1. *Describing computational scores of personality language use (PLU)*

Again, all the computational variables were standardized, maintaining the interpretation of the computational scores consistent across the constructed response task. Fig. 4 illustrates the variability of the PLU indicators within the hierarchical semantic vector subspaces, revealing distinct patterns of language utilization across the different personality traits. The linear relationships between the direct and inverse language indicators were negative and small for openness, null for extraversion, small for neuroticism and conscientiousness, and medium for agreeableness, which serves as further evidence of the relatively orthogonal nature of the direct and inverse language indicators of PLU.

### 3.2. *Further validity evidence of personality language use (PLU) in dimensions and facets of the Big Five questionnaire*

Table 5 presents the model fit results for different levels of invariance in each of the personality traits. Configural invariance was established as the same configuration for both PLU profiles, while structural invariance was established as the same configuration and same structural paths (that is, same language-personality relationships) for both PLU profiles. All the models presented good model fit to the data according to

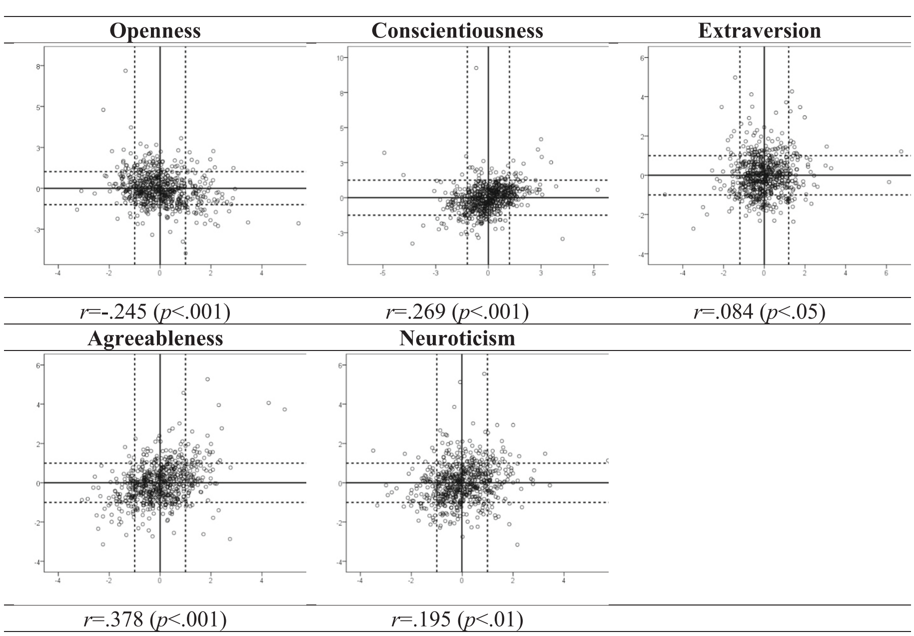


**Fig. 4.** Relationships between direct and inverse language indicators of personality traits in the hierarchical semantic vector subspaces (self-descriptions for a social profile).
*Note*. x-axis: direct indicator. y-axis: inverse indicator. Continuous lines: means. Discontinuous lines: ±1 standard deviation.

**Table 5**
Model fit results (prediction of self-report inventories by language indicators) for different levels of invariance comparing profiles of personality language use (PLU).

| Personality trait | Model invariance | $\chi^2(df)$ | CFI | TLI | RMSEA [90%CI] | SRMSR | Model comparisons | | |
|---|---|---|---|---|---|---|---|---|---|
| | | | | | | | $\Delta\chi^2$ | Δdf | *p*-value |
| Openness | Configural | 626.562 (584) | 0.993 | 0.991 | 0.015 [0.000–0.024] | 0.051 | 28.511 | 9 | <0.001 |
| | Structural | 655.073* (593) | 0.989 | 0.987 | 0.018 [0.005–0.026] | 0.052 | | | |
| Conscientiousness | None | 502.741** (271) | 0.968 | 0.962 | 0.037 [0.032–0.042] | 0.053 | – | – | – |
| Extraversion | Configural | 789.891** (580) | 0.947 | 0.935 | 0.034 [0.028–0.039] | 0.060 | 9.495 | 9 | 0.3929 |
| | Structural | 799.386** (589) | 0.946 | 0.936 | 0.033 [0.027–0.039] | 0.060 | | | |
| Agreeableness | Configural | 749.920** (580) | 0.958 | 0.949 | 0.030 [0.024–0.036] | 0.061 | 68.501 | 9 | <0.001 |
| | Structural | 818.421** (589) | 0.943 | 0.943 | 0.035 [0.029–0.041] | 0.066 | | | |
| Neuroticism | Configural | 886.588** (588) | 0.979 | 0.975 | 0.040 [0.034–0.045] | 0.061 | 34.472 | 9 | <0.001 |
| | Structural | 921.060** (597) | 0.977 | 0.973 | 0.041 [0.036–0.046] | 0.062 | | | |

*Note.* * = $p < .05$. ** = $p < .01$. Robust model comparisons were conducted for DWLS estimator. Models were estimated separately for each personality trait. No invariance models were estimated for contentiousness because of the sample size of one of the latent profiles was smaller than the number of variables.

standard criteria, however, the structural paths of openness, agreeableness, and neuroticism showed statistically significant differences. The structural paths were relatively similar for extraversion, and no invariance could be studied for conscientiousness because of the sample size of one of the latent profiles was smaller than the number of variables. Table 6 presents the estimated structural paths (i.e., language-personality relationships) for the dimensions and facets of the different personality traits of the questionnaire. First, it was found that the personality-language relations are more present in some profiles of PLU than in others. Openness is an exception, as the relations between the different language indicators and its dimensions and facets are similar in both profiles. Second, the language indicators seem to present statistically significant relations with all the dimensions of the Big Five questionnaire, while their relations are more variable in the facets of the self-report inventory. Third, some interesting differences can be seen between the dimensions and the facets of each personality trait, with the PLU of conscientiousness and extraversion only presenting statistically significant relations at the dimension-level, while the PLU of openness presented similar relations with both its dimensions and facets. On the contrary, the PLU of agreeableness and neuroticism showed a significantly larger relation with the facets comparing to their relations with the dimensions of the questionnaire, especially with cordiality and control of emotions, respectively. The results of conscientiousness were conducted for the whole sample, showing a statistically significant structural path between the indirect language indicator and the conscientiousness dimension ($b = 0.067$, $SE = 0.023$, $p < .01$), although it did not present a statistically significant relation with scrupulousness ($b = -0.008$, $SE = 0.040$, $p = .848$) or determination ($b = -0.010$, $SE = 0.028$, $p = .718$). The direct language indicator did not present any statistically significant structural path with conscientiousness ($b = 0.026$, $SE = 0.025$, $p = .310$), scrupulousness ($b = 0.042$, $SE = 0.042$, $p = .322$), or determination ($b = 0.031$, $SE = 0.029$, $p = .287$).

## 4. Discussion

In the present study, we tested a proof of concept by generating low-cost semantic indicators that capture relevant variability of personality traits from language using semantic vector subspaces. Using new developments in VSMs, we generated small vector subspaces whose computational scores (coordinates) were able to capture specific and general semantic meanings associated to the target constructs (e.g., Jorge-Botana et al., 2019; Martínez-Huertas et al., 2022). In this case, the resulting semantic vector subspaces were able to capture semantic properties of language that were common and specific of the target personality traits. In this sense, we have proposed a formal model of the semantics of the PLU of the Big Five personality model (McCrae & Costa Jr., 2008). These preliminary results suggest that these semantic vector subspaces can be used to extract personality trait-relevant semantic properties from language.

We found that there could be different groups or patterns of PLU. The latent profiles found in the present study showed convergent and discriminant validity regarding the target personality trait. This means that there could be specific patterns of self-descriptive language use for the target personality traits. Though the patterns identified here may be a unique finding of this study, we deem it worthwhile to further explore these results, as each profile exhibits vastly different response patterns, thereby enabling us to classify the diverse personality traits estimated by the Big Five personality model. While we cannot generalize these patterns of language use, we do deem the proposal of subspaces as trait classification systems to be a tool of great potential.

In the first study, it was found convergent and discriminant validity between the latent profiles of PLU (i.e., a discrete variable) and the

**Table 6**
Structural path estimates of the structural equation models testing the prediction of self-report inventories by language indicators.

| Personality trait | Variable | Language indicators of profile 1 | | | Language indicators of profile 2 | | |
|---|---|---|---|---|---|---|---|
| | | Direct | Indirect | General | Direct | Indirect | General |
| **Openness** | **Openness** | .226*** (.023) | -.073** (.027) | .068** (.025) | .225** (.075) | .079 (.072) | .181** (.068) |
| | **Cultural openness** | .139*** (.038) | -.077t (.039) | -.057 (.038) | .117 (.091) | -0.058 (.079) | .062 (.084) |
| | **Openness to experience** | .079** (.029) | -.014 (.03) | .086** (.030) | .229* (.096) | .209* (.086) | .102 (.077) |
| **Extraversion** | **Extraversion** | .009 (.030) | -.048 (.031) | .006 (.029) | .157*** (.052) | -0.045 (.038) | .020 (.048) |
| | **Dynamism** | -.041 (.048) | .097t (.050) | .007 (.048) | -0.188t (.100) | -0.047 (.067) | .122 (.068) |
| | **Dominance** | .056 (.039) | -.016 (.041) | .053 (.040) | .026 (.069) | -0.030 (.056) | .056 (.069) |
| **Agreeableness** | **Agreeableness** | .010 (.029) | .058 (.030) | .098** (.030) | -.019 (.045) | -.171** (.052) | .026 (.049) |
| | **Cordiality** | .068 (.055) | -.041 (.052) | -.056 (.055) | -.557*** (.098) | -.772*** (.127) | -.777*** (.127) |
| | **Cooperation** | .025 (.047) | -.037 (.046) | -.033 (.046) | .017 (.070) | .111t (.061) | -.049 (.064) |
| **Neuroticism** | **Neuroticism** | .050t (.027) | -.039 (.028) | .007 (.027) | .338*** (.097) | -.122t (.063) | .520*** (.108) |
| | **Control of emotions** | -.060 (.040) | -.042 (.042) | -.037 (.040) | .811** (.238) | .051 (.120) | .825** (.265) |
| | **Impulse control** | .030 (.041) | .015 (.041) | .044 (.042) | -.190 (.127) | .308** (.100) | -.197 (.141) |

*Note.* $t = p < .10$. * = $p < .05$. ** = $p < .01$. *** = $p < .001$. Language indicators were standardized, and the structural paths can be interpreted as the effect size. Models were estimated separately for each personality trait.

personality traits measured with standardized self-report inventories. In the second study, linear relations were discovered between the PLU indicators and the personality traits measured with standardized self-report inventories, but only for certain latent profiles. In this sense, we also found clear linear relations between the computational language indicators and the personality traits for specific groups of participants. This suggests that there could be individual differences in the way people express their personality through language. In other words, there would be groups of people where linear relations could be expected between their PLU and the usual standardized self-report inventories. On the contrary, it is worth mentioning that the results may be suggesting that most of the participants have a "usual" or "common" PLU whose relations with personality traits are not linear or nonexistent, and may perhaps fall outside of the radar of our language screening. Probably, this could decrease the likelihood of finding significant language-personality relations in other empirical studies, particularly in similar prompted-based self-descriptions. Therefore, it is crucial to continue profiling the individual characteristics that make the personality of participants more capable of being evaluated using properties of language from formal computational models. Additionally, we found evidence in favor of the Brunswik symmetry principle as, for some personality traits, the language-personality relations were observed in the Big Five dimensions with larger effect sizes than in its facets. In the case of openness, we found a similar size of the language-personality relation in both the dimension and its facets. Similarly, the language indicators only had a relevant linear relation with the trait dimension of conscientiousness and extraversion (showing not-statistically significant linear relations with their facets). Conversely, we found larger linear relations between the language indicators and some facets of agreeableness and neuroticism (cordiality and control of emotions, respectively) comparing to the dimensions. These results suggest a complex pattern of results where some Big Five personality traits would present relations with language at the dimension level while others would be related to language at the facet level, suggesting another illustration of the Brunswik symmetry principle.

From a substantive point of view, the current model has some interesting insights for the study of PLU. It was found that people who share similar personality characteristics tends to present a similar self-descriptive semantics. This finding partially supports the well-known *lexical hypothesis* of personality (e.g., Allport & Odbert, 1936; Ashton & Lee, 2005; Cattell, 1943; Fiske, 1949; Goldberg, 1993; John et al., 1988), which is crucial to understand the origins of the Big Five personality model (Goldberg, 1993). Basically, two main postulates have been associated to the *lexical hypothesis* of personality which strengthens the use of formal computational models of language: language captures personality characteristics that are important to groups or societies, and those relevant personality characteristics may be encoded as single words in some cases. These proposals are not new and were initiated more than a century ago by the hand of authors like Galton (1884, seen in Galton, 1949) in his "Measurement of Character" paper. Other classic defenders of the *lexical hypothesis* of personality were the well-known study of Allport and Odbert (1936), and the posterior analyses or dimensionality reduction of that study made by Cattell (1943) and Fiske (1949). While these studies were revolutionary and profoundly influenced the psychological science of personality, we believe that it is possible to go beyond by the hand of formal computational models of language due to their systematicity and capacity of extending the study of semantics to thousands of words significantly easing the Big Five personality screening in texts (e.g., self-descriptions, social network contents, real conversations, etc.). In this sense, as it was stated previously, these PLU indicators could present transparent relations with personality traits only for some participants, which could hone or contextualize the predictions of the *lexical hypothesis* of personality.

From a methodological point of view, the present study illustrates how to formalize the semantics of language, through semantic vector subspaces, to ease the Big Five personality screening in prompted-based self-descriptions. To the best of our knowledge, this is the first extension to measure not-cognitive constructs -like personality traits- using semantic vector subspaces, as the transformation of VSMs were previously used to capture semantic concepts to solve cognitive tasks (e.g., Jorge-Botana et al., 2019; Martínez-Huertas et al., 2022; Martínez-Mingo et al., 2023). However, while the use of computational methods for the analysis of language in relation to personality traits has allowed for a more efficient utilization of information, a fundamental issue remains: the delimitation of the terms used in relation to a specific trait. The proposed method addresses this issue by identifying the use or non-use of language terms related to each personality trait individually, through the delimitation of subspaces. These new vector spaces are defined by dimensions that correspond to each personality trait, forcing any represented term to be signified in relation to that specific trait. This approach not only aligns with the study of language, but also with many other areas of cognitive psychology. Nonetheless, the present study is a proof of concept. The descriptors to construct the direct and inverse vectors in each trait were constrained by the descriptors of the TIPI test. The problem is that there were too few words for the two dimensions and the vector for the new basis for each personality trait was constructed only with two or three words choosing the simplest language of current psychometric models to capture PLU. Nevertheless, there are other "data-driven" alternatives to generate semantic definitions of personality traits. For example, Kjell et al. (2019, 2020) evaluated different psychological constructs using language generated by participants to reach definitions of the target variables. This proposal could be implemented in the context of personality research by asking participants to produce descriptors of, for example, people that are extrovert and introvert, and generating vector semantic subspaces from them.

We found promising results working with limited language indicators generated from the simplest language of the available psychometric models. But it is not necessary to be so pure in applied research. Among other extensions, we could evaluate which words from a set of items generate the best performance. Similarly, these language indicators could be included as input of predictive models (being the orthogonality of semantic vector subspaces one clear advantage). In this sense, further research should extend the use of this methodology in other Large Language Models (LLMs) like, for example, neural network models like RNN-LSTMs and Transformers based LLMs. An advantage of using these LLMs is that descriptions can be represented by means of contextual embeddings, which consider compositional properties of the sentences, comparing to the static embeddings of other models like LSA or *word2vec*. By incorporating such contextual embeddings in our proposal, we could use LLMs without requiring the large costs of generating thousands labeled utterances to train these models for a specific task. This would imply to define personality-based subspaces in the word embedding space. Of course, there is still a long road ahead to get this methodology to work with not orthogonal semantic representations of LLMs (e.g., BERT, ELMo, GPT models, etc.), but it is a research line that we will pursue in the following years.

Regarding to the limitations of this paper, we only evaluated the convergent and discriminant validity of the scores of the semantic vector subspaces, ignoring other evidence of construct validity (e.g., criterion validity). We conceive the convergent and discriminant validity as the first steps to obtain a complete construct validity of these computational scores. Also, we must acknowledge that the convergent validity of the standardized self-report inventories was far from perfect, which is a common concern in personality research (e.g., Pace & Brannick, 2010), and limited our results. Another limitation was the use of undergraduate students as the sample of the study, as some range restrictions could be expected in the measured personality traits. We hypothesize to find more extreme profiles of PLU in representative samples of the population. Also, the present studies were conducted in Spanish, although this methodology is language-independent and applicable to any language. In this sense, we only analyzed two constructed responses which were prompted-based self-descriptions, but these semantic vector subspaces should also be tested in other types of language-based tasks. In any case, we think that this limitation is also a strength of the study because of the selected language-based tasks were relatively different as one consisted of the enumeration of self-descriptive words and the other consisted of the generation of a profile for a social network. Thus, we showed the generalizability of the computational scores to different types of self-descriptive language-based tasks.

### *4.1. Then, where is our proposal positioned in the current scientific panorama?*

We are living great times in the research of LLMs, as there are relevant advances almost every week. LLMs can be accommodated to solve tasks like the prediction/screening of personality from texts. We see two general trends about using semantic vector representations of prompted-based self-descriptions to predict personality in current psychological literature: (1) supervised learning (techniques that require a specific training and large data samples that include the target constructs -labels/scores- to learn to solve the task), and (2) unsupervised learning (those that do not require them). Our proposal is included in the latter.

Supervised learning comprehends data-driven predictive models that can be trained to predict personality from some specific characteristics encoded in vector representations of text responses. These predictive models are tuned to predict a personality trait from a specific task. But this requires relatively large data sets with information of the predictive variables (text vector representations) and the dependent variable (scores of personality traits). Although metanalytic studies support the use of predictive models in this context (Moreno et al., 2021), it is very expensive and a priori not generalizable to new/different types of input because of their specific-task-specialization.

Unsupervised learning includes many techniques that do not take part of a predictive model trained to solve a specific task using a specific input (although this does not exclude the possibility of incorporating them in other predictive models). Here, we consider the existence of "data-driven" and "theory-driven" approaches. There are some interesting "data-driven" proposals like, for example, trying to account for the semantics of constructs by asking participants to produce word norms that define the target variables (e.g., Kjell et al., 2019, 2020; this procedure includes a predictive model based on multiple linear regressions, but the word norms of the constructs are produced by participants). In contrast, "theory-driven" proposals would be guided by a priori information about the constructs. For example, researchers may develop personality-specific dictionaries to define the traits. The *proof of concept* of this paper would be located within those "theory-driven" proposals, as it operationalizes the definition of the personality traits from the available language of psychometric models. The words that we used here to generate the semantic vector subspaces are not perfect semantic definitions of the Big Five personality traits, but we found evidence in favor of this perspective using the simplest available language.

While computational capacity has somewhat alleviated the challenge of sample size, a more pressing issue lies in the need to define distinct and specific learning systems for predicting personality traits from text. This requirement could be mitigated through semantic vector subspaces that enhance interpretability, enabling us to *'open the black box'* and explore the semantic meanings of predictive models. Moreover, the usefulness of semantic vector subspaces is transversal to all these approaches as their small-dimensional vector representations can be used in both unsupervised and supervised learning studies. Thus, we believe that semantic vector subspaces are useful and parsimonious ways to capture (at least, partially) some variability of personality.

## Open practices statement

The semantic vector subspaces for screening PLU in texts, the R code, and an empirical illustration are provided in the OSF project: https://doi.org/10.17605/OSF.IO/26SQY.

## CRediT authorship contribution statement

**José Ángel Martínez-Huertas:** Writing – original draft, Writing – review & editing. **Guillermo Jorge-Botana:** Writing – original draft, Writing – review & editing. **Alejandro Martínez-Mingo:** Writing – review & editing. **José David Moreno:** Writing – review & editing. **Ricardo Olmos:** Writing – review & editing.

## Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## Data availability

The semantic vector subspaces for screening personality language use in texts, the R code, and an empirical illustration are provided in the OSF project

## Acknowledgements

The first author would like to thank the Asociación Española de Metodología de las Ciencias del Comportamiento (AEMCCO) for their support to develop this research project by awarding him with the Grant AEMCCO-Julio Olea to young researchers in 2021.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.paid.2023.112514.

## References

Allport, G. W., & Odbert, H. S. (1936). *Trait-names: A psycho-lexical study*. Psychological Review Company.

Ashton, M. C., & Lee, K. (2005). A defence of the lexical approach to the study of personality structure. *European Journal of Personality, 19*(1), 5–24. https://doi.org/10.1002/per.541

Bermúdez, J. (2001). Cuestionario "big five". In *Adaptación al castellano del cuestionario BFQ de Caprara, Barbaranelli, & Borgogni (1993)*. TEA: Ediciones.

Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: A new approach to personality in a digital world. *Current Opinion in Behavioral Sciences, 18*, 63–68. https://doi.org/10.1016/j.cobeha. 2017.07.017

Cattell, R. B. (1943). The description of personality: Basic traits resolved into clusters. *The Journal of Abnormal and Social Psychology, 38*(4), 476–506. https://doi.org/10.1037/h0054116

Digman, J. M. (1997). Higher-order factors of the big five. *Journal of Personality and Social Psychology, 73*(6), 1246. https://doi.org/10.1037/0022–3514.73.6.1246

Eysenck, H. J. (1994). The big five or the giant three: Criteria for a paradigm. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 37–51). Erlbaum.

Fast, L. A., & Funder, D. C. (2008). Personality as manifest in word use: Correlations with self-report, acquaintance report, and behavior. *Journal of Personality and Social Psychology, 94*(2), 334–346. https://doi.org/10.1037/0022-3514.94.2.334

Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology, 44*(3), 329–344. https://doi.org/10.1037/h0057198

Galton, F. (1949). The measurement of character. In W. Dennis (Ed.), *Readings in general psychology* (pp. 435–444). Prentice-Hall, Inc.. https://doi.org/10.1037/11352-058

Garten, J., Hoover, J., Johnson, K. M., Boghrati, R., Iskiwitch, C., & Dehghani, M. (2018). Dictionaries and distributions: Combining expert knowledge and large scale textual data content analysis. *Behavior Research Methods, 50*(1), 344–361. https://doi.org/10.3758/s13428-017-0875-9

Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist, 48*(1), 26–34. https://doi.org/10.1037/0003-066X.48.1.26

Gosling, S. D., Rentfrow, P. J., & Swann, W. B., Jr. (2003). A very brief measure of the big five personality domains. *Journal of Research in Personality, 37*, 504–528. https://doi.org/10.1016/S0092-6566(03)00046-1

Günther, F., Rinaldi, L., & Marelli, M. (2019). Vector-space models of semantic representation from a cognitive perspective: A discussion of common misconceptions. *Perspectives on Psycho- logical Science, 14*(6), 1006–1033. https://doi.org/10.1177/1745691619861372

Hu, X., Cai, Z., Wiemer-Hastings, P., Graesser, A. C., & McNamara, D. S. (2007). Strengths, limitations, and extensions of LSA. In T. K. Landauer, D. S. Mcnamara,

S. Dennis, & W. Kintsch (Eds.), *Handbook of latent semantic analysis* (pp. 401–426). Routledge. https://doi.org/10.4324/97 80203936399.ch20.

John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality, 2*(3), 171–203. https://doi.org/10.1002/per.2410020302

Jorge-Botana, G., Olmos, R., & Barroso, A. (2013, July). Gallito 2.0: A natural language process- ing tool to support research on discourse. In *Proceedings of the twenty-third annual meeting of the Society for Text and Discourse, Valencia*.

Jorge-Botana, G., Olmos, R., & Luzón, J. M. (2019). Could LSA become a "Bifactor" model? Towards a model with general and group factors. *Expert Systems with Applications, 131*, 71–80. https://doi.org/10.1016/j.eswa.2019.04.055

Jorge-Botana, G., Olmos, R., & Luzón, J. M. (2020). Bridging the theoretical gap between semantic representation models without the pressure of a ranking: Some lessons learnt from LSA. *Cognitive Processing, 21*, 1–21. https://doi.org/10.1007/s10339-019-00934-x

Kjell, O. N. E., Kjell, K., Garcia, D., & Sikström, S. (2019). Semantic measures: Using natural language processing to measure, differentiate, and describe psychological constructs. *Psychological Methods, 24*(1), 92–115. https://doi.org/10.1037/met0000191

Kjell, O. N. E., Kjell, K., Garcia, D., & Sikström, S. (2020). Semantic similarity scales: Using semantic similarity scale to measure depression and worry. In S. Sikström, & D. Garcia (Eds.), *Statistical semantics: Methods and applications* (pp. 53–74). Springer.

Kundu, A., Jain, V., Kumar, S., & Chandra, C. (2015). A journey from normative to behavioral operations in supply chain management: A review using latent semantic analysis. *Expert Systems with Applications, 42*(2), 796–809. https://doi.org/10.1016/j.eswa.2014.08.035

Kwantes, P. J., Derbentseva, N., Lam, Q., Vartanian, O., & Marmurek, H. H. (2016). Assessing the Big Five personality traits with latent semantic analysis. *Personality and Individual Differences, 102*, 229–233. https://doi.org/10.1016/j.paid.2016.07.010

Landauer, T. K., & Dumais, S. (1997). A solution to Plato's problem: The latent semantic analysis theory of the acquisition, induction, and representation of knowledge. *Psychological Review, 104*, 211–240. https://doi.org/10.1037/0033-295X.104.2.211

Martínez-Huertas, J. A., Olmos, R., Jorge-Botana, G., & León, J. A. (2022). Distilling vector space model scores for the assessment of constructed responses with bifactor inbuilt rubric method and latent variables. *Behavior Research Methods, 54*, 2579–2601. https://doi.org/10.3758/s13428-021-01764-6

Martínez-Huertas, J. A., Olmos, R., & León, J. A. (2021). Enhancing topic-detection in computerized assessments of constructed responses with distributional models of language. *Expert Systems with Applications, 185*(115621), 1–12. https://doi.org/10.1016/j.eswa.2021.115621

Martínez-Mingo, A., Jorge-Botana, G., Martínez-Huertas, J. A., & Olmos, R. (2023). Quantum projections on conceptual subspaces. *Cognitive Systems Research, 82* (101154), 1–13. https://doi.org/10.1016/j.cogsys.2023.101154

McCrae, R. R., & Costa, P. T., Jr. (2008). The five-factor theory of personality. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 159–181). Guilford.

McCrae, R. R., & Sutin, A. R. (2018). A five-factor theory perspective on causal analysis. *European Journal of Personality, 32*(3), 151–166. https://doi.org/10.1002/per.2134

McNamara, D. S. (2011). Computational methods to extract meaning from text and advance theories of human cognition. *Topics in Cognitive Science, 3*(1), 3–17. https://doi.org/10.1111/j.1756-8765.2010. 01117.x

Moreno, J. D., Martínez-Huertas, J. A., Olmos, R., Jorge-Botana, G., & Botella, J. (2021). *Can personality traits be measured by analyzing written language?* A meta-analytic study on computational methods: Personality and Individual Differences. https://doi.org/10.1016/j.paid.2021. 110818

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus User's Guide* (8th ed.). Muthén & Muthén.

Nesselroade, J. R., & McArdle, J. J. (1997). On the mismatching of levels of abstraction in mathe- matical-statistical model fitting. In M. D. Franzen, & H. W. Reese (Eds.), *Biological and neuropsychological mechanisms: Life-span Developmental Psychology* (pp. 23–49). Erlbaum.

Pace, V. L., & Brannick, M. T. (2010). How similar are personality scales of the "same" construct? A meta-analytic investigation. *Personality and Individual Differences, 49*(7), 669–676. https://doi.org/10.1016/j.paid.2010.06.014

Pennebaker, J. W., Francis, M. E., & Booth, R. J. (2001). *Linguistic inquiry and word count: LIWC 2001*. Erlbaum.

Pennebaker, J. W., & Graybeal, A. (2001). Patterns of natural language use: Disclosure, personality, and social integration. *Current Directions in Psychological Science, 10*(3), 90–93. https://doi.org/10.1111/1467-8721.00123

Pennebaker, J. W., & King, L. A. (1999). Linguistic styles: Language use as an individual difference. *Journal of Personality and Social Psychology, 77*(6), 1296–1312. https://doi.org/10.1037/0022-3514. 77.6.1296

Rosseel, Y. (2011). Lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Vigil-Colet, A., Morales-Vives, F., Camps, E., Tous, J., & Lorenzo-Seva, U. (2013). Development and validation of the overall personality assessment scale (OPERAS). *Psicothema, 25*(1), 100–106. https://doi.org/10.7334/psicothema2011.411

Visinescu, L. L., & Evangelopoulos, N. (2014). Orthogonal rotations in latent semantic analysis: An empirical study. *Decision Support Systems, 62*, 131–143. https://doi.org/10.1016/j.dss.2014.03.010

Wittmann, W. W. (1988). Multivariate reliability theory: Principles of symmetry and successful validation strategies. In J. R. Nesselroade, & R. B. Cattell (Eds.), *Handbook of multivariate experimental psychology* (pp. 505–560). Plenum Press.